DOCUMENT RESUME

ED 137 672 CG 011 258

AUTHOR Prediger, Dale J.; Hanson, Gary R.
TITLE Evidence Related to Issues of Sex Bias in Interest Inventories.
PUB DATE [76]
NOTE 28p.; Paper presented at the Annual Conference of the American Psychological Association (84th, Washington, D.C. September 3-7, 1976)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Career Planning; Goal Orientation; Interest Research; Interest Tests; Occupational Guidance; Personality Tests; Predictive Ability (Testing); Raw Scores; *Sex Discrimination; Symposia; *Test Bias; *Vocational Interests

ABSTRACT

This symposium examines recent research on sex restrictiveness as a potential indicator of sex bias. Two specific questions are addressed: (1) Can useful non-sex-restrictive raw score interest scales be constructed? Research results are reported for more than 10,000 persons in six samples including 9th graders, 11th graders, college-bound students, college sophomores, and adults. (2) Are sex-restrictive interest score reports more valid than non-restrictive reports? Data for two longitudinal validity studies (Ns=1,400 and 1,000) and two concurrent validity studies (Ns=1,400 and 10,500) are summarized. Results of this research, together with research on the construct validity of sex-restrictive and non-restrictive reporting procedures, indicate that sex restrictiveness in interest score reports is not a necessary concomitant of validity; hence it may be indicative of sex bias. (Author/JLL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED137672

EVIDENCE RELATED TO ISSUES OF SEX BIAS IN INTEREST INVENTORIES[1]

Dale J. Prediger and Gary R. Hanson

Research and Development Division
American College Testing Program

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Many interest inventories in common use (e.g., the Kuder General Interest Survey, Ohio Vocational Interest Survey, Self-Directed Search, ACT Interest Inventory) report scores for general types of career interests (e.g., mechanical, clerical, social, etc.). Even the "Strong," which is normally thought of as focusing on specific occupations, provides scores for the 6 Holland interest types and 23 additional general interest areas. This paper addresses issues related to sex restrictiveness and sex bias in the score reports provided by inventories that assess general types of career interests.

Importance of Problem

When reports of general interests are based on the raw scores obtained by counselees, as in Holland's Self-Directed Search (SDS), they are typically sex-restrictive. That is, males and females are referred in large numbers to occupations traditional to their sex (Gottfredson, Holland, & Gottfredson, 1975; Holland, 1972; Prediger & Hanson, 1974). For example, the distributions of SDS high-point codes (highest scores) for "diverse samples of 2,169 high school boys [and] 2,447 high school girls" (Gottfredson et al., 1975, p. 139) are as follows:

Holland type	Girls	Boys	Occupational category associated with type
S	67%	20%	Education and social welfare occupations
A	13	8	Artistic, musical, and literary occupations
C	11	3	Office and clerical occupations
R	1	40	Skilled trades, technical, and some service occupations
I	8	23	Scientific and some technical occupations
E	1	6	Managerial and sales occupations

As can be seen, over 90% of the high school girls receive their highest SDS raw scores for S, A, or C (social, artistic, office, clerical, etc.) occupations. Only about 10% score highest on I, R, or E (scientific, trades/technical, managerial, etc.) occupations. In contrast, 70% of the boys receive their highest scores for these latter occupations.

Because nearly 1,000,000 persons have taken the SDS since its introduction (Consulting Psychologist Press, 1976), its potential impact on the career plans of American youth is great. If raw scores were used to report results for the ACT

CG 011258

---

[1] Paper providing basis for summary presentations at the 1976 American Psychological Association National Convention symposium with same title. Presentation titles covered by paper are as follows:

Prediger: "Is Sex Restrictiveness a Necessary Concomitant of Validity?"
Hanson: "Criterion-related Validity of Sex-restrictive Reports of Interests"
Schussel: "Sex Differences in Interest Inventory Item Responses: Necessity or Artifact?"

Interest Inventory (ACT-IV), each year some 400,000 college-bound girls (about 85% of those tested) would receive their highest scores for S, A, and C occupations. Nearly 60% of the college-bound boys would score highest on the I, R, and E scales. We cannot justify that consequence of using raw scores on psychometric, theoretical, or ethical grounds. Hence, we report standard scores for the ACT-IV. Because Cole and Hanson (1975) have demonstrated that interest score reports based on combined-sex norms tend to be as sex-restrictive as those based on raw scores, we base the standard scores on same-sex norms.

When same-sex norms are used to report interest inventory results, similar proportions of males and females receive their highest scores for a given occupational area (e.g., natural science, business management, social work). In this respect, score reports based on same-sex norms are sex-balanced and those based on raw scores or combined-sex norms are sex-restrictive. If same-sex norms were used in reporting SDS results, there would be few if any differences in the proportion of males and females referred to a given category of occupations. Thus, there are dramatic differences in the career guidance provided by sex-restrictive and sex-balanced score reports. The remainder of this paper examines the appropriateness of sex-restrictive reporting procedures.

Issue 1: How Does One Determine Sex Bias?

Recent definitions of sex bias (e.g., AMEG Commission on Sex Bias in Measurement, 1973; Diamond, 1975) focus on the sex-limiting (sex-restrictive) effects of interest inventories. For example, the working definition used in developing the NIE guidelines for assessing sex bias in interest inventories (Diamond, 1975) is as follows: "Within the context of career guidance, sex bias is defined as any factor that might influence a person to limit [restrict] . . . his or her consideration of a career solely on the basis of gender" (p. xxiii). If the consideration of career options is limited solely on the basis of gender, then the distribution of career options suggested to males and females by an interest inventory will differ. Stated another way, the scores reported by the inventory will suggest one set of occupations to most males and a different set to most females, as in the SDS example. Sex-restrictive career guidance will be provided. But is a sex-restrictive interest inventory necessarily sex biased?

Prediger and Hanson (1974) maintain that sex-restrictiveness is a characteristic of interest inventories that should be studied separately. According to their definition of sex bias, a sex-restrictive interest inventory is not sex biased if it can be shown that sex-restrictiveness is a necessary concomitant of validity. That is, the mere presence of sex-restrictiveness, however socially significant, is not sufficient to establish sex bias. Sex bias can be established only if the publisher of a sex-restrictive inventory can not demonstrate that sex restrictiveness is a necessary concomitant of validity. Thus, the definition follows principles underlying Equal Employment Opportunity Commission Guidelines (1970). The burden of proof, however, is on the publisher, not the user.

Most publishers of interest inventories will have scores for general samples of persons for whom the inventories are intended. If they do not, they should. By applying to these scores the interpretive procedures they advocate, publishers can determine the distribution of career options suggested to males and females. Counselors who care about sex bias in career counseling can demand that publishers provide these distributions. When sex-restrictiveness is evident, counselors can also demand evidence that it is a necessary concomitant of validity. We believe that these steps define a practical, empirical approach to detecting sex bias in interest inventories, especially if followed in conjunction with the NIE Guidelines. We hope counselors, as individuals or in association with others, will give them a try.

Holland (1975, 1976) has suggested quite a different approach to the determination of sex bias in interest inventories. According to Holland (1975), "an inventory is unbiased when its experimental effects on female and male respondents are similar and of about the same magnitude--that is, when a person acquires more vocational options, becomes more certain, or learns more about himself (herself) and the world of work" (p. xxvii). On the surface, this definition is quite appealing. Sex bias is determined by the effects of an inventory on counselees. However, Holland's definition ignores the nature of the effects--the nature of the career guidance provided by inventories. According to the definition, an interest inventory that provides highly stereotypic career options to males and females would not be sex biased if males and females in similar numbers pursued these options, became more certain of their choices, etc. It matters not what the career options are so long as the effects on males and females "are similar and of about the same magnitude."

Because it ignores the career options suggested by an interest inventory, Holland's definition of sex bias appears to miss the mark. Evidence that counselees pursue the career options suggested by an interest inventory does not mean the inventory is sex fair. The suggestions themselves may be sex biased. If an interest inventory has an effect on counselees, it is primarily through the career options it suggests. Thus, it would appear that in determining sex restrictiveness or bias in interest inventories, the focus should be on the career options suggested to counselees.

Issue 2: Should Predictions of Occupational Preference and Entry Be Used in Validating Interest Inventories?

The validity of a measuring instrument depends on the purposes for which it is used. Hence, before studying validity, one must ask "Validity for what?" Interest inventories are commonly used to suggest occupational options for counselees to consider. Yet, the validity of an inventory is sometimes determined by its ability to predict future occupational preferences or occupational entry. As Berdie (1970) has noted, few counselors are interested in predicting whether a counselee will enter (or prefer) occupation A or occupation B. Hence, validity data for this use of interest inventories may not be relevant to the intended use. Furthermore, interest inventories with high validity for predicting occupational preference or entry may produce harmful side effects. Some of the reasons are discussed below and an alternative approach to validation is proposed.

Validation Using a "Will-Prefer-or-Enter" Criterion

When predicting the occupations persons will prefer or enter, one must take into account the nature of occupational preference and employment distributions. That is, if an interest inventory is to provide accurate predictions of eventual employment, the predictions must accurately reflect the size of each occupational criterion group. To the degree that group membership predictions depart from group base rates, the inventory's predictive accuracy will be lowered.

Interest inventories which predict that males and females will enter or prefer occupations in the same proportions as in the past should have high "hit rates" under this approach to validation. For a multitude of reasons (e.g., social expectations, local labor market needs, the contingencies of life), people will continue to state preferences for and enter traditional occupations. Unfortunately, the occupational preference and employment distributions of males and females are highly divergent (Gottfredson, Holland, & Gottfredson, 1975; Prediger, Roth, & Noeth, 1974). Since the predictions used in validation studies are based on the scores received by counselees, the occupational options suggested to counselees will reflect the same

divergencies as the predictions, an unfortunate side effect of this approach to validation. The SDS distributions presented earlier show what can happen.

Validation Using a "Should-Consider" Criterion

The alternative approach to the use of occupational preference and membership as criteria in validating interest inventories assumes that the purpose of interest inventories is to identify career options for counselees to consider rather than to predict the occupations counselees will prefer or enter. To achieve the former objective, an interest inventory must assess the correspondence between a counselee's interests and the interests associated with various occupational groups--regardless of the group base rates. If Cindy's interests are compatible with engineering, one would suggest that she, and others like her, consider engineering even if this lowers the accuracy with which occupational entry is predicted. The emphasis is on "should consider," not "will enter or prefer." In the course of career exploration, Cindy should find out that there are relatively few women engineers but that the situation is changing. Those facts may play a role in her career decisions; they should not influence her interest scores.

Studies following this approach to interest inventory validation will treat occupational criterion groups (e.g., preference groups) as if they were of equal size. One would expect an interest inventory to suggest engineering to a large proportion of criterion group members in engineering, nursing to a large proportion of nurses, retail sales to retail sales clerks, horse shoeing to horse shoers, and so on for each of the criterion groups available. It does not matter that there are relatively few horse shoers in comparison to retail sales clerks. The question asked in the validation analysis is "What proportion of the members of each criterion group would have been asked by this interest inventory to look into their occupation?" Stated differently, the question is "What is the hit rate for each criterion group?" A high hit rate depends on an inventory's ability to differentiate the criterion groups and thus minimize the misassignment of members of each of the groups.

In this approach to validation, an interest inventory does not have to suggest retail sales to more counselees than horse shoeing because there are more retail sales clerks than horse shoers. "Predictions" are simply based on whichever criterion group a person resembles most. No premium is placed on providing interest score distributions that parallel preference or employment distributions. The proposed validation strategy recognizes that, for a number of very practical reasons, many persons may not enter the occupations suggested ("predicted") by an interest inventory.

How Choice of Criterion Affects Career Guidance

The following example may bring into sharper focus the practical implications of the two approaches to validation. Suppose that in a society built on the caste system, an interest inventory were designed to have high validity in predicting occupational entry. The inventory would suggest few if any occupations that were not traditional for a person's caste. To do otherwise would lower its validity. On the other hand, suppose an inventory were designed to identify occupational options compatible with a person's interests--regardless of the proscriptions of this society. Such an inventory might suggest many occupations not traditional for members of castes. As a result, it would be a poor predictor of occupational entry. Yet, it might do an excellent job of determining occupational compatibility. Even in a time of social change, the score reports for such an inventory might be unsettling. But they could provide beneficial information, both to the individual and to the society.

Although useful in some types of research, interest inventories designed to predict which persons will prefer or enter a given occupation present special problems for career counseling. In effect, the rationale underlying such inventories says "Cindy may have interests like an engineer and Mike may have interests like a nurse. But few females or males are likely to enter those nontraditional occupations. So let your predictions (and career guidance) take into account the relative numbers of males and females who have entered various occupations in the past. In the long run, you'll obtain a higher hit rate and your inventory will appear to be more valid." When used in career counseling, such inventories will reinforce the society's occupational sex-role stereotypes and thus further institutionalize the channeling. At first glance, these inventories may appear to have higher validity than inventories designed to report occupational options compatible with a person's interests. But this may be true only if one's purpose in assessing interests is to predict the occupations counselees will enter (or prefer). Counselors concerned with sex-bias in interest measurement should determine the main purposes for which they use interest inventories.

Prediger and Cole (1975) provide an extended discussion of this topic as it applies to career counseling and nontraditional occupations for males and females. Prediger (1976) discusses implications for validation procedures.

Issue 3: Must Interest Score Reports be Sex-Restrictive in Order to be Valid?

Although it is possible to eliminate sex-restrictiveness from interest inventory score reports through the use of same-sex norms, Holland and his coworkers maintain that sex-restrictive score reports simply reflect the effects of the sex-role socialization process, that interest score reports must be sex-restrictive in order to be valid, and that score reports which are not sex-restrictive will have lowered validity (Gottfredson et al., 1975; Gottfredson & Holland, 1975; Holland, 1974, 1975). For example, Gottfredson et al. (1975) argue that the interest score distributions of men and women should correspond to the occupational distributions of men and women. That is, the proportion of women (or men) referred to an occupational field by an interest inventory should correspond to the proportion of women (or men) currently employed in that field. Holland and his coworkers demonstrate that sex-restrictive raw scores for Holland types approximate this criterion far better than sex-balanced normed scores and thus presumably support the validity of raw scores. However, Prediger and Cole (1975) question the theoretical and measurement basis for using employment distributions as criteria for judging the adequacy of interest score distributions.

In another study, Gottfredson and Holland (1975) compare the criterion-related validity of SDS raw scores and normed scores (same-sex norms) in predicting the vocational preferences of college students. Even if one agrees that the purpose of interest inventories is to predict vocational preference, the results obtained by Gottfredson and Holland appear to be ambigious. Raw scores and normed scores were equally effective predictors for males. Although raw scores provided more accurate predictions than normed scores for females, both sets of predictions were less accurate than predictions based on criterion group size (i.e., the base rates). These latter predictions ignore SDS scores. Finally, current occupational preference was a better predictor of future preference than the SDS for both males and females. This latter finding is not surprising since current preferences consistently predict future preferences better than do interest inventories (Gottfredson & Holland, 1975; Holland, 1973). However, it implies that interest inventories may be of little practical value to counselors who want to predict the future vocational preferences of their counselees.

Criterion-related Validity

The validity of interest inventories for identifying occupations counselees should consider has been investigated in several studies. (See discussion of this approach to validation under Issue 2.) In a reanalysis of the Gottfredson-Holland data for college women, Prediger (1976a) found that the validities of raw scores and normed scores did not differ. Hit rates averaged approximately 40% for both types of scores. ACT-IV data for a national sample of college seniors (5,517 men and 5,061 women) in the same 18 majors also show that the criterion-related validity of normed scores is equal to or greater than that of raw scores (Prediger & Hanson, Note 3). For males, the average hit rate for high-point code predictions of college majors grouped by Holland type was 55% for standard scores (same-sex norms) and 53% for raw scores. For females, the hit rate was 51% for standard scores and 44% for raw scores. When centour scores based on discriminant functions were used, the hit rates for normed scores and raw scores were essentially the same.

The criterion-related validity of several different interest inventory reporting procedures was examined using 2-year and 5-year longitudinal data from two studies (Hanson, Noeth, & Prediger, Note 2). The first study involved 1,072 students who began community college in the fall of 1970 and who were subsequently surveyed during the spring and summer of 1975. Students were categorized into criterion groups corresponding to Holland types on the basis of current employment or employment plans. Most of the individuals in the latter category (20% of total group) were still enrolled in college. High-point codes based on each individual's 1970 results from the Vocational Interest Profile (VIP), an alternate form of the ACT-IV, were matched against criterion group membership. A "hit" was tallied if the student's highest score in 1970 corresponded to the student's criterion group. The unweighted average hit rates were higher for scores based on same-sex norms than for raw scores for both men (38% vs. 33%) and women (44% vs. 36%). For women, same-sex norms produced higher hit rates than combined-sex norms and opposite-sex norms. For men, they were superior to opposite-sex norms and equal in effectiveness to combined-sex norms.

The second longitudinal study involved 1,443 college-bound students who took the ACT-IV in the fall of 1972 and were subsequently surveyed during the spring of 1975. ACT-IV high-point codes were matched against college major categorized by Holland type. For males, the unweighted average hit rates were similar for reports based on same-sex norms (43%), raw scores (42%), and combined-sex norms (40%). For females, same-sex and combined-sex norms produced identical hit rates (39%), both superior to raw scores (33%). Data for opposite-sex norms were not examined.

The results of these two longitudinal studies, the Prediger-Hanson study, and the Gottfredson-Holland study indicate that, for common criteria, the criterion-related validity of scores based on same-sex norms is equal to or greater than that of the other reporting procedures. This finding is especially important since raw scores, combined-sex norms, and opposite-sex norms all provide highly stereotypic career suggestions to males and females whereas same-sex norms do not.

Finally, Hanson and Rayman (in press) recently reported the results of a criterion-related validity study comparing sex-restrictive and sex-balanced raw score scales. (The construction of sex-balanced raw score scales is discussed under Issue 4.) Six criterion groups for males (N = 502) and five for females (N = 876) were formed by categorizing the current vocational preferences of a national sample of college-bound students tested in April 1974. When centour scores based on discriminant functions were used to obtain group membership predictions, the average hit rates for the two types of raw score scales (sex-restrictive and sex-balanced) were essentially the same for both males and females.

Construct Validity

Three comparisons of the construct validity of sex-restrictive and sex-balanced interest score reports have also been completed recently. The first study (Prediger & Hanson, 1976) involved more than 39,000 men and women pursuing the same 104 occupations. Differences in the raw score profiles of men and women pursuing the same occupations were substantial, systematic, and stereotypic. When same-sex norms were used, the profiles of men and women were much more similar--in accordance with expectations based on Holland's theory of careers (Holland, 1976; Prediger & Hanson, 1976). In a closely related study involving the sample of 10,500 college seniors described earlier, ACT-IV profiles for males and females in the same 18 majors were also found to be more similar when same-sex norms were used (Hanson, Note 1). In the third study (Prediger, in press-a, in press-b) bearing on the construct validity of sex-restrictive and sex-balanced score reports, normed scores produced personality pattern distributions more in line with expectations based on Holland's theory than did raw scores. Results were replicated across seven samples of high school students, college students, and adults, and four different interest inventories. More than 18,000 males and 20,000 females were involved.

Tentative Resolution of Issue 3

Results from each of the eight studies cited above indicate that the validity of interest inventories is not lowered through the use of sex-balanced score reports. In several instances, it is increased. A recent study by Lamb (1975) also indicates that sex-balanced score reports are appropriate for use with males and females in various minority groups, and several studies show that persons in various college majors (ACT, 1972; Hanson, 1974), occupational preference groups (ACT, 1974), and occupations (Prediger & Hanson, 1976) obtain sensible score profiles when sex-balanced scales are used. Considered as a whole, the evidence strongly suggests that sex-restrictiveness is not a necessary concomitant of validity. Thus, sex-restrictive interest score reports may well be sex biased.

Issue 4: Can Useful Sex-Balanced Raw Score Scales Be Constructed?

Clearly, it is possible to construct sex-balanced interest inventory scales through the use of same-sex norms. Such norms automatically provide similar score distributions for males and females and, hence, similar distributions of career options (e.g., see Gottfredson et al., 1975; and Prediger & Hanson, 1974). Although same-sex norms have traditionally been used in reporting career interests, they are currently the subject of considerable misunderstanding. Holland and his coworkers (Gottfredson et al., 1975), for example, refer to same-sex norms as "norms which treat men and women differently" (p. 140). Indeed, different norms tables are used for males and females in order to avoid sex-restrictive score reports. Unfortunately, "treating men and women the same" (using raw scores or combined-sex norms) will not eliminate sex-restrictiveness in typical interest inventories. The sex-restrictive career options suggested to males and females by raw scores and combined sex-norms were discussed earlier.

One way to avoid the misunderstandings sometimes associated with the use of same-sex norms is to eliminate sex-restrictiveness in the interest inventory raw scores themselves, and thus make same-sex norms unnecessary. Because the responses to individual items provide the basis for raw scores, attempts to construct "sex-balanced raw score scales" (i.e., raw score scales on which males and females obtain approximately equal scores) mainly involve field testing potentially sex-balanced items. Items in existing scales that elicit large male-female response differences can then be replaced by new, sex-balanced items with the appropriate psychometric characteristics.

Results Obtained by Rewording Current Items

It is clear from the work of Holland and Gottfredson that minor revisions in the wording of items typically used in interest inventories are not likely to have an appreciable effect on the raw score distributions of males and females. For example, Gottfredson (1976) found essentially no differences in the responses of 94 high school girls to 4 "sexist" and 4 "neutral" occupational titles (e.g., policeman and police officer). On this basis, Gottfredson concluded that his approach to modifying current items will not change overall raw scores to any practical degree. Holland and Gottfredson (1976) examined the effect of replacing 12 of 22 SDS Realistic Scale items rarely endorsed by women and items foreign to their experience with items thought to be sex-neutral (e.g., prefer to take a shop or sewing course). They found that the Realistic Scale scores of women were significantly and substantially increased. However, the distribution of SDS high-point codes did not change, probably because of methodological limitations in the study. As the authors point out, "women usually get such high scores on the social, artistic, and conventional scales that only a great change in their realistic scale scores would change their SDS code" (p. 225). Holland and Gottfredson did not try to revise the other scales. In addition, several Realistic Scale items likely to produce sex differences were not revised. Finally, only 2 of the 12 items written for the study appeared to measure the intended construct. These two studies show that attempts to modify a few of the items typically included in interest inventories will probably have little effect on sex-restrictiveness.

Results Obtained Through Substantial Item Revisions

No attempt was made in either of the above studies to write and pretest items endorsed in equal proportions by men and women or to develop raw score scales providing similar distributions of career options for men and women. However, the fact that many interest inventory items in common use elicit similar responses from men and women suggested that it may be possible to construct an interest inventory consisting entirely of sex-balanced items. A recent study by Rayman (1976) demonstrated that interest inventory items on which there are only minor sex differences can be written and that raw score scales developed from such items possess psychometric characteristics (e.g., scale homogeneity) similar to those of scales containing items that elicit large sex differences. Rayman constructed potentially sex-balanced items for each of Holland's six types. Following a pretest with 220 high school seniors, the items were administered to a national sample of 1,902 college-bound students along with the ACT-IV. In contrast to ACT-IV raw scores, Rayman's Unisex Interest Inventory (UNI-II) raw scores did not exhibit the large sex differences typically found. Nevertheless, the pattern of interscale correlations for the UNI-II corresponded to the hexogonal configuration expected for Holland types and correlations between the UNI-II scales and the ACT-IV scales showed the appropriate convergent and discriminant validity. Finally, as noted in the previous discussion of the study by Hanson and Rayman (in press), the criterion-related validities of centour scores based on discriminant functions were similar for the sex-balanced UNI-II raw scores and the sex-restrictive ACT-IV raw scores.

Development and Characteristics of the Uni-Sex ACT Interest Inventory

Because the use of sex-balanced interest inventory items is in accordance with the National Institute of Education (NIE) Guidelines for Assessment of Sex Bias and Sex Fairness in Career Interest Inventories (Diamond, 1975), and because this procedure makes it possible to use a combined-sex norms table without restricting the occupational options suggested to either sex (as would typically be the case; Cole & Hanson, 1975), ACT decided in the spring of 1974 to explore the development of a form of the ACT Interest Inventory (ACT-IV) based on sex-balanced items. Subsequent to the work

of Rayman (1976), researchers at ACT conducted six studies of the characteristics of potentially sex-balanced interest items. More than 10,000 persons including 9th graders, 11th graders, college-bound students, college sophomores, and adults were involved in the studies. Eight scales were developed--six scales designed to assess the same six interest dimensions assessed by the ACT-IV (Hanson, 1974) and two bipolar scales designed to assess data/ideas and people/things work-related activity preferences (Prediger, 1976). The new Uni-Sex Interest Inventory (UNIACT-IV) scales provide highly similar distributions of career options to males and females. Internal consistency estimates of reliability for the 15-item scales range from .85 to .92 with a median of .87. Correlations with the six parallel ACT-IV scales range from .76 to .86 with a median of .80. The pattern of intercorrelations for the UNIACT-IV scales closely approximates the hexagonal configuration expected on the basis of Holland's theory of careers (Holland, 1973). Correlations between the two 30-item bipolar scales and factor scores representing the data/ideas and people/things dimensions are .84 and .83, respectively. All of these data are based on holdout samples not involved in item selection and scale construction. An overview of UNIACT-IV development and characteristics is provided in the appendix to this paper.

Tentative Resolution of Issue 4

Data for the UNIACT-IV, together with the research results reported by Rayman (1976) and Hanson and Rayman (in press), indicate that sex-balanced raw score scales can be constructed for interest inventories assessing basic types of interests and that these scales possess several desirable psychometric characteristics. Most important, they suggest similar career options to males and females. As noted under Issue 3, the appropriateness of this result is well-supported by research on interest score reports based on same-sex norms. In effect, when one constructs sex-balanced scales, one norms at the item level.

Interest inventories in common use contain sex-balanced items along with items that elicit stereotypic responses from males and females. However, the latter items appear to be an unfortunate legacy from an earlier era of measurement and guidance. Although interest scales based on such items may prove helpful in studying the effects of sex-role socialization on self-reported preferences for various occupations and activities, their use in suggesting career options to counselees is open to serious question.

Reference Notes

1. Hanson, G. R. The career interests of men and women in the same college majors. Manuscript in preparation, 1976.

2. Hanson, G. R., Noeth, R. J., & Prediger, D. J. The validity of diverse procedures for reporting interest scores: An analysis of longitudinal data. Manuscript in preparation, 1976.

3. Prediger, D. J., & Hanson, G. R. Some consequences of using raw score reports of vocational interests. Manuscript submitted for publication, 1976.

References

The American College Testing Program. Handbook for the Career Planning Program. (1972 ed.) Iowa City, Iowa: Author, 1972.

The American College Testing Program. Handbook for the Career Planning Program, Grades 8-11. Boston: Houghton Mifflin, 1974.

Association for Measurement and Evaluation in Guidance Commission on Sex Bias in Measurement. Report on sex bias in interest measurement. Measurement and Evaluation in Guidance, 1973, 6, 171-177.

Berdie, R. F. Vocational choice studies. Personnel and Guidance Journal, 1970, 48, 604-605.

Cole, N. S., & Hanson, G. R. Impact of interest inventories on career choice. In E. E. Diamond (Ed.), Issues of sex bias and sex fairness in career interest measurement (National Institute of Education Report). Washington, D.C.: U.S. Government Printing Office, 1975.

Consulting Psychologist Press. 1976 Catalog. Palo Alto, Calif.: Author, 1976.

Diamond, E. E. (Ed.). Issues of sex bias and sex fairness in career interest measurement (National Institute of Education Report). Washington, D.C.: U.S. Government Printing Office, 1975.

Equal Employment Opportunity Commission. Guidelines on employment selection procedures. Federal Register, 1 August 1970, 35, 12,333.

Gottfredson, G. D. A note on sexist wording in interest measurement. Measurement and Evaluation in Guidance, 1976, 8, 221-223.

Gottfredson, G. D., & Holland, J. L. Vocational choices of men and women: A comparison of predictors from the Self-Directed Search. Journal of Counseling Psychology, 1975, 22, 28-34.

Gottfredson, G. D., Holland, J. L., & Gottfredson, L. S. The relation of vocational aspirations and assessments to employment reality. Journal of Vocational Behavior, 1975, 7, 135-148.

Hanson, G. R. Assessing the interests of college youth: Summary of research and applications (ACT Research Report No. 67). Iowa City, Iowa: The American College Testing Program, 1974.

Hanson, G. R., & Rayman, J. Validity of sex-balanced interest inventory scales. Journal of Vocational Behavior, in press.

Holland, J. L. Professional manual for the Self-Directed Search. Palo Alto, Calif.: Consulting Psychologists Press, 1972.

Holland, J. L. Making vocational choices: A theory of careers. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Holland, J. L. Some guidelines for reducing systematic biases in the delivery of vocational services. Measurement and Evaluation in Guidance, 1974, 6, 210-218.

Holland, J. L. The use and evaluation of interest inventories and simulations. In E. E. Diamond (Ed.), Issues of sex bias and sex fairness in career interest measurement (National Institute of Education Report). Washington, D.C.: U.S. Government Printing Office, 1975.

Holland, J. L. The virtues of the SDS and its associated typology: A response to Prediger and Hanson. Journal of Vocational Behavior, 1976, 8, 349-358.

Holland, J. L., & Gottfredson, G. D. Sex differences, item revisions, validity, and Self-Directed Search. Measurement and Evaluation in Guidance, 1976, 8, 224-228.

Lamb, R. R. Concurrent validity of the American College Testing Interest Inventory for minority group members. (Doctoral dissertation, University of Iowa, 1974). Dissertation Abstracts International, 1975, 35, 7, 4161-A. (University Microfilm No. 75-1216)

Prediger, D. J. Alternatives for validating interest inventories against group membership criteria (ACT Research Report No. 76). Iowa City, Iowa: The American College Testing Program, 1976. (a) (Also Applied Psychological Measurement, in press.)

Prediger, D. J. A world of work map for career exploration. Vocational Guidance Quarterly, 1976, 24, 198-208. (b)

Prediger, D. J. The viability of Holland's consistency construct and raw score assessments of personality. Measurement and Evaluation in Guidance, in press-a.

Prediger, D. J. Do raw scores deserve a D minus? A reply to Holland. Measurement and Evaluation in Guidance, in press-b.

Prediger, D. J., & Cole, N. S. Sex-role socialization and employment realities: Implications for vocational interest measures. Journal of Vocational Behavior, 1975, 7, 239-251.

Prediger, D. J., & Hanson, G. R. The distinction between sex restrictiveness and sex bias in interest inventories. Measurement and Evaluation in Guidance, 1974, 7, 96-104.

Prediger, D. J., & Hanson, G. R. Holland's theory of careers applied to women and men: Analysis of implicit assumptions. Journal of Vocational Behavior, 1976, 8, 167-184.

Prediger, D. J., Roth, J. D., & Noeth, R. J. Career development of youth: A nationwide study. Personnel and Guidance Journal, 1974, 53, 97-104.

Rayman J. Sex and the single interest inventory: The empirical validation of sex-balanced interest inventory items. Journal of Counseling Psychology, 1976, 23, 239-246.

APPENDIX[1]

THE UNI-SEX ACT INTEREST INVENTORY (UNIACT-IV): OVERVIEW OF DEVELOPMENT AND CHARACTERISTICS

Gary R. Hanson, Dale J. Prediger, and Robert H. Schussel

Research & Development Division

American College Testing Program

Overview of Development

Campbell (1974) suggests that a difference of 15% or more between two groups is a substantial and meaningful difference in the response to an interest inventory item. For purposes of a working definition during the construction of UNIACT-IV, an item was considered to be "sex-balanced" if the difference in the percentage of "like" responses between men and women was 10% or less, a somewhat more conservative approach than that taken by Campbell. The same working definition was used throughout all the item tryouts, although a few items in final pool did not meet this criterion.

A summary of the samples used in item tryouts is provided in Table 1.

Insert Table 1 about here

Hereafter, these samples will be referred to by letter designation (e.g., Sample A, B, C, etc.). More than 200 potentially sex-balanced items were administered to these samples. Many of the items underwent repeated tryouts.

For all samples except Sample B, the potentially sex-balanced items were administered concurrently with the ACT Interest Inventory (ACT-IV) (Hanson, 1974) or one of its alternate forms. For Sample B, the items were administered within 8 weeks after students completed the ACT-IV. There were three more or distinct stages of item selection. The first involved Samples A and B; the

---

[1]Paper distributed at the 1976 American Psychological Association National Convention symposium entitled "Evidence Related to Issues of Sex Bias in Interest Inventories."

second involved Samples C and D; and the third involved Samples E and F. At each stage, item selection and refinement were determined by the degree of balance in the percentage of "like" responses for males and females and by the correlation of items with the various ACT-IV scales. The balance in the percentage of "dislike" responses between males and females was also noted. Items which showed a 10% difference or less in the percentage of "like" responses and which correlated with the appropriate scales on the ACT-IV were retained for subsequent tryouts. For Samples D, E, and F, potentially sex-balanced items were also correlated with data/ideas and people/things factor scores based on the theoretical definitions of these two bipolar dimensions. (See section on factor structure that follows.) Items which correlated with the appropriate factor score and which were sex-balanced were retained for possible inclusion on the final scales.

Samples B, C, E, and F were subdivided into an item analysis sample and a holdout (cross-validation) sample. Items selected from the item analysis samples were scored on the appropriate UNIACT-IV preliminary scales and the resulting scale means and frequency distributions were checked for overall sex balance using the holdout samples. Scale intercorrelations were also examined. Data from the item analysis subsamples of Samples E and F were used in making final refinements in the UNIACT-IV scales.

UNIACT-IV Scale Characteristics

Unless otherwise noted, all analyses reported in this section were conducted on a sample of 1,851 students (914 males and 937 females) obtained by combining holdout groups from Samples E and F with an additional 177 11th grade students not involved in previous analyses.

Degree of Sex Balance

Table 2 presents UNIACT-IV means, standard deviations, and the percent

---

Insert Table 2 about here

---

overlap for the scale distributions of males and females. The first six scales, all 15-item scales corresponding to the six types of interests ("personal orientations") proposed in Holland's theory of careers (Holland, 1973), have names identical to the six scales in the ACT-IV. The Social Service and Technical Scales, which correspond to Holland types traditionally exhibiting large sex differences (social and realistic), overlap more than 80%. Although perfect sex balance has not been achieved, data on the overlap of ACT-IV raw scores for males and females, also presented in Table 2, indicate that sex balance was improved substantially for the scales assessing Holland's six types. Furthermore, a comparison of the interest scores of the analysis sample and the scores of a nationally representative sample indicates a greater difference in the interests of males and females in the analysis sample. Thus, the percent overlap for males and females in general will be higher than is shown in Table 2.

The distributions for males and females on the Data/Ideas Scale overlap 95% which indicates excellent sex balance. The percentage overlap for males and females on the Things/People Scale (76%) is somewhat lower but still substantial. For the most part, males and females as groups obtain similar though not identical scores on these two bipolar scales.

The data/ideas and things/people scores provided by UNIACT-IV can be translated into a region on the World of Work Map (Prediger, 1976) used in the ACT Assessment Program (the "ACT") to encourage students to explore occupations in various world of work "regions." Hence, the distribution of region scores obtained by males and females provide a more direct comparison of the career suggestions males and females receive from UNIACT-IV than does scale overlap. These distributions are presented in Table 3 for males and females in the nationally representative UNIACT-IV norm group. Although there is some tendency for males

---

Insert Table 3 about here

---

to be referred more often to "things" occupations (regions 6 and 7) and females more often to "ideas" occupations (regions 9 and 10), the differences are not great, especially when considered in light of the highly divergent career suggestions provided by popular reporting procedures.

Reliability

Means, standard deviations, internal consistency estimates of reliability, and intercorrelations for UNIACT-IV scales are shown in Table 4. For the six scales

---

Insert Table 4 about here

---

assessing Holland types, the coefficient alpha reliability estimates range from .85 to .92, with a median value of .87, a rather high value considering the relatively short length (15 items) of the scales. Although test-retest reliabilities for the scales have not yet been determined, stability coefficients for similar 15-item scales in the ACT-IV ranged from .80 to .89 for retesting after 8 weeks. The median was .85 (Hanson, 1974). The median 9-week stability coefficient for the 15-item scales in the Vocational Interest Profile, also a forerunner of UNIACT-IV, was .80; the range was .73 to .85 (ACT, 1974).

Because items comprising the two bipolar interest scales were weighted in opposite directions, coefficient alpha estimates of reliability could not easily be calculated. Instead, split-half reliability estimates were obtained for the two 30-item scales. The split-half estimates for the Data/Ideas and Things/People Scales were .75 and .82, respectively, figures which suggest that the two interest constructs are measured with considerable reliability, especially since the contribution of response set variance to reliability is minimized on bipolar scales. The scale reliability appears to be quite adequate for reports based on World of Work Map regions (in contrast to exact scores).

Factor Structure

Figure 1 depicts the theoretical relationship between the six UNIACT-IV scales and the data/ideas and things/people dimensions (Prediger, 1976). For example, the

theoretical model suggests that the Data/Ideas Scale should have high positive correlations with the Business Detail and Business Contact Scales and high negative correlations with the Science and Creative Art Scales. The correlations (reported in Table 4) for the six scales assessing Holland types and the two bipolar scales are in general agreement with expectations. In addition, the pattern of correlations is similar to the pattern of loadings for the theoretical factors described below. Finally, the two scales are essentially independent (r = .03), as expected.

Insert Figure 1 about here

Another source of evidence related to the factor structure of the UNIACT-IV is the degree to which the six UNIACT-IV scales conform to the expected two-dimensional circular configuration proposed by Holland (1973) and illustrated in Figure 1. An analysis of spatial configuration (Cole & Cole, 1971) was used to show the inter-relationships among the six UNIACT-IV scales in two-dimensional space. In the configuration resulting from this analysis, scales which correlate most highly are located closest together in space while scales with lower correlations are located farther apart.

The results of the configural analysis are shown in Figure 2. The two dimensions on which the six scales are located account for 57% of the variation

---
Insert Figure 2 about here
---

of the scale locations in six-dimensional space, a value highly similar to findings for other inventories (Cole & Hanson, 1971; Cole, 1973; Hanson, 1974). The relative locations of the six scales conform to theoretical expectations to a high degree. The circular configuration proposed by Holland is reproduced with nearly equal distance between scales, and the scales are arranged in the expected order. Because there is a high degree of similarity between the configuration that was obtained and the theoretical configuration shown in Figure 1, it is evident that the six UNIACT-IV scales are interrelated in the expected manner and that the

underlying scale structure is in accordance with the data/ideas and things/people dimensions also assessed by UNIACT-IV.

The two right-most columns in Table 5 show the correlations of the six UNIACT-IV Holland-type scales with theoretical factors representing the data/ideas and things/people dimensions. These factors were extracted from the intercorrelation matrix for the six scales by using theoretical expectations to define, a priori, the desired scale loadings. Results indicate that the factors are the primary interest dimensions assessed by the six scales; they are highly similar for males and females; they correspond closely to the factors obtained from principal components analyses of the same data; and they account for about 60% of the scale variance when the effects of response set are removed. (Response set, which is irrelevant to assessment of type of interest, is commonly present in interest inventories having a "like," "dislike," etc. response mode.)

As shown in Table 4, correlations between the UNIACT-IV bipolar scales and the corresponding theoretical factors are .84 and .83 for the Data/Ideas and Things/People Scales, respectively. The correlations are quite high, in light of scale reliabilities, thus indicating that the two scales provide adequate measures of the theoretical dimensions.

Relationship with ACT-IV and VIP Scales

Because the UNIACT-IV, the ACT-IV, and the CPP 8-12 Vocational Interest Profile (VIP) were constructed to assess the same basic types of interests, correlations for corresponding scales should be substantial. However, UNIACT-IV was designed to assess preferences for basic categories of career-related activities independent of sex-role expectations. Items on UNIACT-IV involve activities associated with basic types of interests. The items are endorsed in similar proportions by males and females. Such items should be much less subject to stereotypic reactions to the object or nature of an activity than many of the items on the ACT-IV and other interest inventories (e.g., take dictation, repair a hot rod, operate a power shovel). To the degree that sex-role expectations affect the

scores on interest inventories, irrelevant variance is introduced, unless of course one's purpose is to access stereotypic attitudes toward various types of activities, occupational titles, etc. Hence, one would expect only moderate correlations between UNIACT-IV scores and the scores from instruments not containing sex-balanced scales.

Concurrent correlations between the six pairs of same-named scales on the UNIACT-IV and the ACT-IV ranged from .76 to .86 for the holdout subgroup of Sample E; the median correlation was .80. Correlations for same-named scales on the UNIACT-IV and VIP, based on the holdout subgroup of Sample F, ranged from .72 to .81 with a median value of .74. These correlations represent a relatively high degree of relationship given the reliability levels of these relatively short scales and differences in the sex-balance of item content. The overall level of correlation appears to be sufficiently high to permit the generalization of validity data for the ACT-IV (Hanson, 1974) and the VIP (ACT, 1974) to the UNIACT-IV.

References

American College Testing Program. Career Planning Program, Grades 8-11, Handbook. Boston, Houghton Mifflin, 1974.

Campbell, D. P. Manual for the Strong-Campbell Interest Inventory. Stanford, Calif.: Stanford University Press 1974.

Cole, N. S. On measuring the vocational interests of women. Journal of Counseling Psychology, 1973, 20, 105-112.

Cole, N. S., & Cole, J. W. L. An analysis of spatial configuration and its application to research in higher education (ACT Research Report No. 35). Iowa City, Iowa: The American College Testing Program, 1970.

Cole, N. S., & Hanson, G. R. An analysis of the structure of vocational interests. Journal of Counseling Psychology, 1971, 18, 478-486.

Dunnette, M. D. Personnel selection and placement. Belmont, Calif.: Wadsworth, 1966.

Hanson, G. R. Assessing the interests of college youth: Summary of research and applications (ACT Research Report No. 67). Iowa City, Iowa: The American College Testing Program, 1974.

Holland, J. L. Making vocational choices: A theory of careers. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Prediger, D. J. A world of work map for career exploration. Vocational Guidance Quarterly, 1976, 24, 198-208.

Table 1

Description of UNIACT-IV Item Tryout Samples

Sample	N	Percent women	Educational level	Nature of sample
A	1,825	63	Grade 14	A sample of college sophomores who responded to a follow-up survey of ACT-IV national norm group (Hanson, 1974).
B				
Item analysis	1,191	59	Mostly grade 12	A nationally representative sample of ACT-tested college-bound students. Sample was obtained by taking every 65th student from the October 1975 national test date. Sample was randomly divided into item analysis and holdout groups.
Holdout	1,191	58		
C				
Item analysis	1,123	49	Grade 9	Sample obtained from 10 rural, small city, and suburban high schools in Iowa and Missouri. Sample was randomly divided into item analysis and holdout groups.
Holdout	725	50		
D	1,250	29	Community college adults	Students 25 years and older enrolled in degree-oriented programs in 10 community colleges in 7 states.
E				
Item analysis	1,031	43	Grade 11	Eleven schools selected from the 200 schools that participated in the national norming of the Career Planning Program, Grades 8-12 (ACT, 1974). Schools were in rural, medium city, large city, and suburban locations in 11 states.
Holdout	1,031	45		
F				
Item analysis	201	50	Grade 11	Five schools chosen in a manner similar to Sample E.
Holdout	643	57		
Total	10,211			

Table 2

UNIACT-IV Means, Standard Deviations, and Percent Overlap for Males and Females

UNIACT-IV scales	Males (N=914)		Females (N=937)		Percent overlap[a]	
	$\bar{X}$	SD	$\bar{X}$	SD	UNIACT-IV	ACT-IV[b]
Science	1.92	.58	1.77	.56	90	91
Creative Arts	1.97	.50	2.04	.52	94	78
Social Service	2.11	.44	2.28	.40	84	61
Business Contact	1.98	.45	1.94	.42	96	90
Business Detail	1.85	.49	1.89	.51	97	93
Technical	1.90	.44	1.70	.42	82	59
Data/Ideas	13.98	5.19	14.58	5.36	95	c
Things/People	14.40	4.26	11.89	3.85	76	c

Note. As explained in the text, the differences between the interest scores of males and females in this sample are more pronounced than for a nationally representative sample.

[a]Based on Dunnette's (1966) table for Tilton's measure of overlap.

[b]Overlap data for ACT-IV raw scores were obtained for the 1,031 students in the holdout group of Sample E and are provided for comparative purposes only. ACT-IV scales were not designed to be sex-balanced and raw scores are not reported.

[c]These scales were not included on the ACT-IV.

Table 3

Distributions of World of Work Map Regions

Suggested to Males and Females

World of work region	Males (N = 462)	Females (N = 468)
1	8.0	6.4
2	8.4	7.5
3	8.2	9.0
4	8.4	7.7
5	7.4	8.5
6	10.0	6.8
7	7.1	6.4
8	8.4	6.8
9	4.3	6.0
10	6.1	10.3
11	6.9	8.1
12	7.1	6.2
Region ?	9.5	10.3

Note. Distributions based on the 10% national sample of 11th grade students in the UNIACT-IV norm group.

Table 4

UNIACT-IV Scale Intercorrelations, Means, Standard Deviations, and Internal Consistency Estimates of Reliability

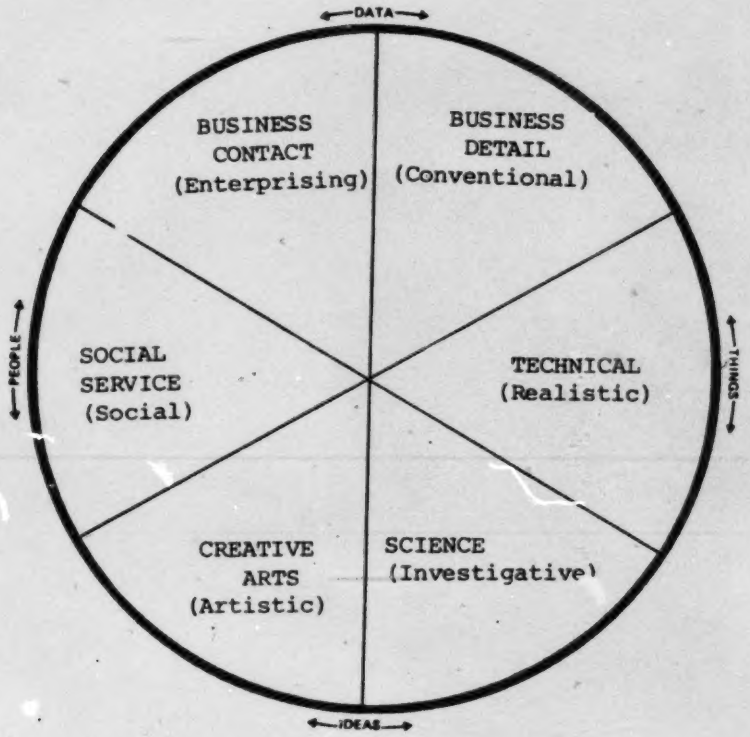
	Science	Creative Arts	Social Service	Business Contact	Business Detail	Technical	Bipolar scales		Theoretical factors	
							Data/Ideas	Things/People	Data/Ideas	Things/Peopl
Science	--	31	32	16	14	27	-60	02	-48	19
Creative Arts	31	--	42	31	06	27	-34	-19	-45	-36
Social Service	32	42	--	55	28	21	-05	-59	04	-60
Business Contact	16	31	55	--	57	30	34	-29	52	-34
Business Detail	14	06	28	57	--	33	53	06	65	17
Technical	27	27	21	30	33	--	00	49	04	50
Data/Ideas	-60	-34	-05	34	53	00	--	03	84	-02
Things/People	02	-19	-59	-29	06	49	03	--	-03	83
D/I factor	-48	-45	04	52	65	04	84	-03	--	00
T/P factor	19	-36	-60	-34	17	50	-02	83	00	--
$\bar{X}$	1.84	2.00	2.20	1.97	1.87	1.80	14.29	13.13	.00	.00
SD	.58	.51	.43	.44	.50	.44	5.29	4.24	.98	.98
Coefficient alpha	.92	.89	.85	.85	.90	.85	.75[a]	.82[a]		

Note. Sample based on 1,851 high school juniors, as described in introduction to section on UNIACT-IV scale characteristics. The same correlational data are reported above and below the diagonal to facilitate examination of the pattern of scale intercorrelations. Decimals are omitted.

[a]Split-half reliabilities.

12

Figure 1. Relationship between UNIACT-IV Scales, Holland Types, and the Data/Ideas, People/Things Work Task Dimensions



Note: UNIACT-IV scale titles are shown in capital letters, Holland types related to the titles are shown in parentheses.

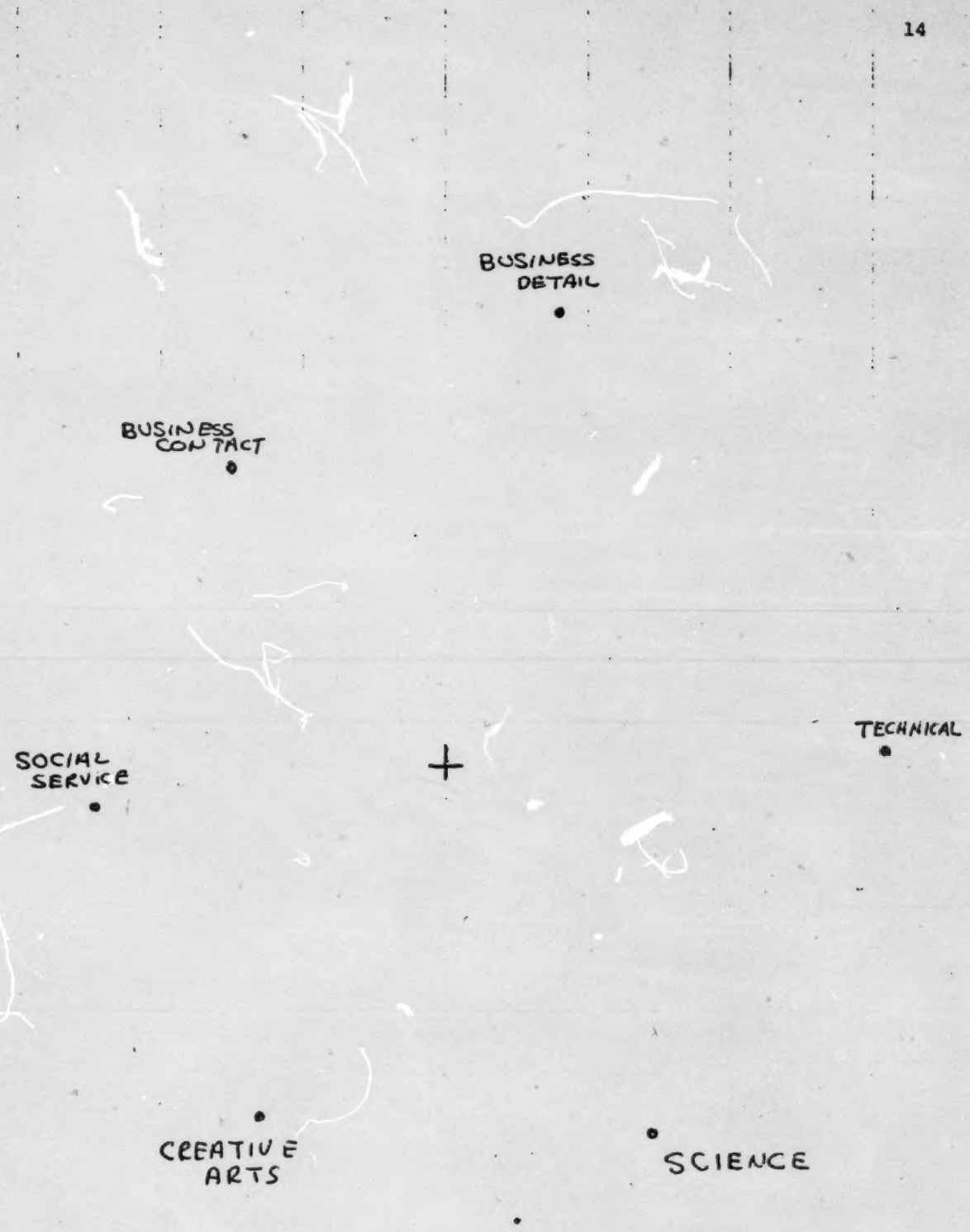


Figure 2. Analysis of Spatial Configuration of the Six UNIACT-IV Scales